

# Low-redshift Lyman limit systems as diagnostics of cosmological inflows and outflows

Zachary Hafen,<sup>1★</sup> Claude-André Faucher-Giguère,<sup>1</sup> Daniel Anglés-Alcázar,<sup>1</sup>  
Dušan Kereš,<sup>2</sup> Robert Feldmann,<sup>3,4</sup> T. K. Chan,<sup>2</sup> Eliot Quataert,<sup>3</sup>  
Norman Murray<sup>5</sup> and Philip F. Hopkins<sup>6</sup>

<sup>1</sup>*Department of Physics and Astronomy and Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA*

<sup>2</sup>*Department of Physics, Center for Astrophysics and Space Sciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 9209, USA*

<sup>3</sup>*Department of Astronomy and Theoretical Astrophysics Center, University of California, Berkeley, CA 94720-3411, USA*

<sup>4</sup>*Institute for Computational Science, University of Zurich, Zurich CH-8057, Switzerland*

<sup>5</sup>*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St George Street, ON M5S 3H8, Canada*

<sup>6</sup>*TAPIR, Mailcode 350-17, California Institute of Technology, Pasadena, CA 91125, USA*

Accepted 2017 April 14. Received 2017 March 27; in original form 2016 August 19

## ABSTRACT

We use cosmological hydrodynamic simulations with stellar feedback from the FIRE (Feedback In Realistic Environments) project to study the physical nature of Lyman limit systems (LLSs) at  $z \leq 1$ . At these low redshifts, LLSs are closely associated with dense gas structures surrounding galaxies, such as galactic winds, dwarf satellites and cool inflows from the intergalactic medium. Our analysis is based on 14 zoom-in simulations covering the halo mass range  $M_h \approx 10^9\text{--}10^{13} M_\odot$  at  $z = 0$ , which we convolve with the dark matter halo mass function to produce cosmological statistics. We find that the majority of cosmologically selected LLSs are associated with haloes in the mass range  $10^{10} \lesssim M_h \lesssim 10^{12} M_\odot$ . The incidence and H I column density distribution of simulated absorbers with columns in the range  $10^{16.2} \leq N_{\text{HI}} \leq 2 \times 10^{20} \text{ cm}^{-2}$  are consistent with observations. High-velocity outflows (with radial velocity exceeding the halo circular velocity by a factor of  $\gtrsim 2$ ) tend to have higher metallicities ( $[X/H] \sim -0.5$ ) while very low metallicity ( $[X/H] < -2$ ) LLSs are typically associated with gas infalling from the intergalactic medium. However, most LLSs occupy an intermediate region in metallicity-radial velocity space, for which there is no clear trend between metallicity and radial kinematics. The overall simulated LLS metallicity distribution has a mean (standard deviation)  $[X/H] = -0.9$  (0.4) and does not show significant evidence for bimodality, in contrast to recent observational studies, but consistent with LLSs arising from haloes with a broad range of masses and metallicities.

**Key words:** galaxies: evolution – galaxies: formation – galaxies: haloes – intergalactic medium – quasars: absorption lines – cosmology: theory.

## 1 INTRODUCTION

The circum-galactic medium (CGM) plays a vital role in determining the evolution of galaxies. Accretion from the intergalactic medium (IGM) passing through the CGM provides the fuel to sustain star formation in galaxies over a Hubble time (e.g. Kereš et al. 2005; Prochaska & Wolfe 2009; Bauermeister, Blitz & Ma 2010). At the same time, galactic feedback processes driven by stars and

active galactic nuclei are observed to expel gas out of galaxies and into the CGM as outflows (e.g. Heckman et al. 2000; Shapley et al. 2003; Martin 2005; Weiner et al. 2009; Steidel et al. 2010; Greene et al. 2011; Jones, Stark & Ellis 2012; Ciccone et al. 2014; Rubin et al. 2014). Cosmological models of galaxy formation furthermore show that strong outflows are necessary to regulate galaxy growth (e.g. Kereš et al. 2009; Anglés-Alcázar et al. 2014; Hopkins et al. 2014; Somerville & Davé 2015) and explain the metal enrichment of the IGM (Aguirre et al. 2001; Oppenheimer & Davé 2006; Booth et al. 2012). At low redshift, Lyman limit systems (LLSs; loosely defined as systems optically thick at the Lyman limit) trace

\* E-mail: [zhafen@u.northwestern.edu](mailto:zhafen@u.northwestern.edu)

overdense structures in the haloes of galaxies and are thus useful observational probes of inflows and outflows (e.g. Ribaud et al. 2011b; Tripp et al. 2011).

In the past few years, the observational census of  $z \leq 1$  LLSs has been greatly improved by new analyses of *Hubble Space Telescope* (*HST*) observations. Ribaud et al. (2011a) measured the incidence of LLS and column density distribution of H I absorbers, using archival data from the Faint Object Spectrograph and Space Telescope Imaging Spectrograph instruments. Thanks to the installation of the Cosmic Origins Spectrograph on *HST*, the increased sample of low-redshift HI-selected LLSs makes it now possible to measure the unbiased metallicity distribution of LLSs. Lehner et al. (2013) measured this metallicity distribution and found evidence for a metallicity bimodality. This measurement included all systems with neutral hydrogen column density in the range  $16.2 \leq \log N_{\text{H I}} \leq 18.5$ . Here and in the rest of this paper, logarithms are to the base 10,  $N_{\text{H I}}$  is in units of  $\text{cm}^{-2}$  and  $M_{\text{h}}$  is in units of  $M_{\odot}$ . Lehner et al. (2013) tentatively identified the metal-rich branch as likely tracing galactic winds, recycled outflows and gas tidally stripped from galaxies. They also noted that the metal-poor branch has properties consistent with the cold accretion streams predicted by cosmological simulations (e.g. Kereš et al. 2005, 2009; Dekel & Birnboim 2006; Brooks et al. 2009; Dekel et al. 2009; Faucher-Giguère, Kereš & Ma 2011; van de Voort et al. 2011). In an analysis that doubles the Lehner et al. absorber sample, Wotta et al. (2016) find broadly consistent evidence for a metallicity bimodality, but note that the bimodality primarily arises from the (more numerous) partial LLSs with column in the range  $16.2 \leq \log N_{\text{H I}} \leq 17.2$ . Quiret et al. (2016) reported hints of a bimodality in the metallicity distribution of super LLSs (SLLSs) in the range  $19 \leq \log N_{\text{H I}} \leq 20.3$  (which they refer to as sub damped Ly  $\alpha$  absorbers) at  $z < 1.25$ , but with lower statistical significance and with no evidence of a metallicity bimodality at higher redshift. Fox et al. (2013) found evidence that metal-rich low- $z$  LLSs tend to show higher O VI columns and broader O VI profiles than metal-poor LLSs.

At the same time, the incidence and metallicity properties of high-redshift LLSs have become increasingly well constrained by ground-based observations (Fumagalli, O’Meara & Prochaska 2011; Fumagalli et al. 2013; Lehner, O’Meara & Fox 2014; Cooper et al. 2015; Prochaska et al. 2015; Fumagalli, O’Meara & Prochaska 2016). Despite this tremendous observational progress, forward modelling of IGM absorbers using cosmological simulations remains critical to understanding the physical nature of LLSs, since spectroscopy only provides direct information on the line-of-sight (LOS) structure of absorbers.

Overall, cosmological simulations have proved very successful at reproducing the observed column density distribution of H I absorbers (Hernquist et al. 1996; Altay et al. 2011; McQuinn, Oh & Faucher-Giguère 2011; Altay et al. 2013; Rahmati et al. 2013; Gurvich, Burkhart & Bird 2016), demonstrating how this technique can be used to inform the physical nature of absorption systems. In general, the agreement between cosmological simulations and observations is best for the Lyman- $\alpha$  forest (e.g. Bird et al. 2013), which is relatively little affected by feedback processes from galaxies (e.g. Kollmeier et al. 2006). On the other hand, simulations predict that a large fraction of circum-galactic LLSs trace galactic winds (Fumagalli et al. 2014; Faucher-Giguère et al. 2015; Rahmati et al. 2015; Faucher-Giguère et al. 2016; Liang, Kravtsov & Agertz 2016), and the distribution of LLSs is therefore sensitive to feedback processes. Most of the other circum-galactic LLSs arise in streams of infalling gas (e.g. Faucher-Giguère & Kereš 2011;

Fumagalli et al. 2011; Goerdt et al. 2012; van de Voort et al. 2012; Faucher-Giguère et al. 2015) or on the outskirts of galaxies. To date, however, most simulation analyses have focused on the properties of LLSs at high redshift (see also Kohler & Gnedin 2007).

Our goal in this paper is to study the properties of simulated LLSs at  $z \leq 1$ , with an eye towards interpreting recent *HST* measurements and developing inflow–outflow diagnostics. We use a set of cosmological zoom-in simulations from the FIRE project (‘Feedback In Realistic Environments’),<sup>1</sup> which implement models for stellar feedback on the scale of star-forming regions. These simulations have been shown to correctly reproduce the star formation histories of galaxies below  $\sim L^*$  (Hopkins et al. 2014), mass–metallicity relations (Ma et al. 2016) at all redshifts where observations are available, and LLS covering fractions in the haloes of  $z \sim 2$  galaxies (Faucher-Giguère et al. 2015; Faucher-Giguère et al. 2016). In these simulations, galaxy-scale outflows are self-consistently generated from the local injection of feedback momentum and energy on small scales (Muratov et al. 2015; Muratov et al. 2017). These successes of the FIRE simulations make them particularly well suited to address the physical nature of LLSs.

For this paper, we define partial LLSs as systems in the range  $16.2 \leq \log N_{\text{H I}} \leq 17.2$ , LLSs as systems in the range  $17.2 \leq \log N_{\text{H I}} \leq 19$ , and SLLSs as systems in the range  $19 \leq \log N_{\text{H I}} \leq 20.3$ . Since partial LLSs and LLSs trace similar structures in the CGM, we will often group them together and refer to them simply as LLSs. Throughout, we assume a standard flat  $\Lambda$  cold dark matter cosmology with  $h \approx 0.7$ ,  $\Omega_{\text{M}} \approx 0.27$  and  $\Omega_{\text{b}} \approx 0.046$ , consistent with the latest observational constraints (e.g. Planck Collaboration XIII 2015).

The structure of this paper is as follows. In Section 2, we describe in more detail the simulations used in our analysis and our methodology to extract observational statistics. In Section 3, we convolve our zoom-in simulations results with the dark matter halo mass function to predict cosmological statistics, including the incidence of LLSs and the H I column density distribution of CGM absorbers. We quantify relationships between LLS column density, metallicity and radial velocity relative to central galaxies in Section 4. In that section, we also compute the overall metallicity distribution of low-redshift LLSs predicted by our simulations. We discuss our results and conclude in Section 5. The Appendix summarizes convergence tests done to validate our analysis.

## 2 METHODOLOGY

### 2.1 Simulation sample

Our analysis is based on cosmological zoom-in (e.g. Porter 1985; Katz & White 1993) simulations from the FIRE project. As in previous FIRE papers, all smooth particle hydrodynamics (SPH) simulations presented in this paper were run with exactly the same code as the original FIRE simulations presented in Hopkins et al. (2014). Therefore, we summarize the key points here only. Our CGM analysis methods are similar to the ones used in Faucher-Giguère et al. (2015) and more details can be found in that paper.

Our simulations were run with the P-SPH (Hopkins 2013) hydrodynamics solver implemented in the GIZMO code (Hopkins 2015). P-SPH uses a pressure–entropy formulation of the (SPH) equations that eliminates the artificial surface tension at contact discontinuities found in traditional density-based SPH formulations (e.g. Agertz

<sup>1</sup> Project website: <http://fire.northwestern.edu>

**Table 1.** Parameters of the new simulations.

Name	$M_h(z=0)$ ( $M_\odot$ )	$M_*(z=0)$ ( $M_\odot$ )	$R_{\text{vir}}(z=0)$ (kpc)	$m_b$ ( $M_\odot$ )	$\epsilon_b$ (pc)	$m_{\text{dm}}$ ( $M_\odot$ )	$\epsilon_{\text{dm}}$ (pc)
m11.4a	2.6e11	6.2e9	180	3.3e4	9	1.7e5	140
m11.9a	8.4e11	3.0e10	250	3.4e4	9	1.7e5	140
MFz0_A2	1.0e13	2.7e11	630	3.0e5	9	1.4e6	140

$M_h(z=0)$  and  $M_*(z=0)$  are the dark matter and stellar masses inside the present-day virial radius,  $R_{\text{vir}}(z=0)$ , defined as in Bryan & Norman (1998).  $m_b$  and  $m_{\text{dm}}$  are the baryonic and dark matter particle masses. The simulations use adaptive gravitational softening lengths for the gas but fixed softening lengths for the dark matter.  $\epsilon_b$  is the minimum force softening length for the gas and  $\epsilon_{\text{dm}}$  is Plummer-equivalent gravitational softening length for the dark matter.

et al. 2007) and resolves the major historical differences between SPH and grid-based codes. The gravity solver in GIZMO is a modified version of the GADGET-3 gravity algorithm (Springel 2005), which implements the adaptive softening method of Price & Monaghan (2007) (which we use for the gas) and a modified softening kernel that represents the exact solution for the potential of the SPH smoothing kernel following Barnes (2012).

In our simulations, gas is allowed to cool to  $\sim 10$  K through atomic and molecular line emission, in addition to the standard processes described by Katz, Weinberg & Hernquist (1996). The FIRE simulations are designed to resolve the Toomre mass/length of gas within galaxies, and therefore the most massive giant molecular clouds in which most star formation occurs in typical galaxies (e.g. Murray 2011). In the simulations, star formation occurs only in dense, locally self-gravitating and molecular/self-shielding gas. Where all these criteria are met, the gas is converted into stars on a local free-fall time. The density threshold for star formation is set  $n_{\text{H}} > 50 \text{ cm}^{-3}$  in most of our simulations, though in practice the self-gravity criterion is generally the most stringent. Stellar feedback is modelled by implementing energy, momentum, mass, and metal return from radiation pressure, photoionization, photoelectric heating, supernovae and stellar winds following the STARBURST99 stellar population synthesis model (Leitherer et al. 1999). We explicitly follow the chemical abundances of nine metal species (C, N, O, Ne, Mg, Si, S, Ca and Fe). During the hydrodynamic simulations, ionization balance of all elements is computed using the cosmic ultraviolet (UV) background model of Faucher-Giguère et al. (2009) and we apply an on-the-fly correction for dense, self-shielded gas based on a local Jeans-length approximation.

We analyse a sample of 14 simulations run to  $z=0$  and spanning the halo mass range  $M_h(z=0) \sim 10^9\text{--}10^{13} M_\odot$ . From previous work, our sample includes the seven main simulations in Hopkins et al. (2014) (m09, m10, m11, m12v, m12q, m12i and m13) and four  $M_h(z=0) \sim 10^{10}\text{--}10^{11} M_\odot$  simulations from Chan et al. (2015) (m10h1297, m10h1146, m10h573 and m11h383). We ran three new simulations (m11.4a, m11.9a and MFz0\_A2) with  $M_h(z=0) \sim 10^{11}\text{--}10^{13} M_\odot$  to better populate that halo mass range for our analysis. Simulation MFz0\_A2 follows the same main halo as the MFz2\_A2 MassiveFIRE simulation (Feldmann et al. 2016), but include a larger high-resolution region and is evolved to  $z=0$ . The simulation parameters for the three new runs are detailed in Table 1. With the exception of m13 and MFz0\_A2, all simulations in our sample have a gas particle mass  $m_b \leq 5 \times 10^4 M_\odot$ , minimum adaptive baryonic smoothing/force softening length  $\epsilon_b \leq 10$  proper pc, dark matter particle mass  $m_{\text{dm}} \leq 5 \times 10^5 M_\odot$ , and dark matter force softening length  $\epsilon_{\text{dm}} = 150$  proper pc. For dwarf galaxies in the sample, the resolution parameters are significantly better. For example, the m09 and m10 runs have a gas particle mass  $m_b \approx 260 M_\odot$  and correspondingly smaller minimum softening lengths.

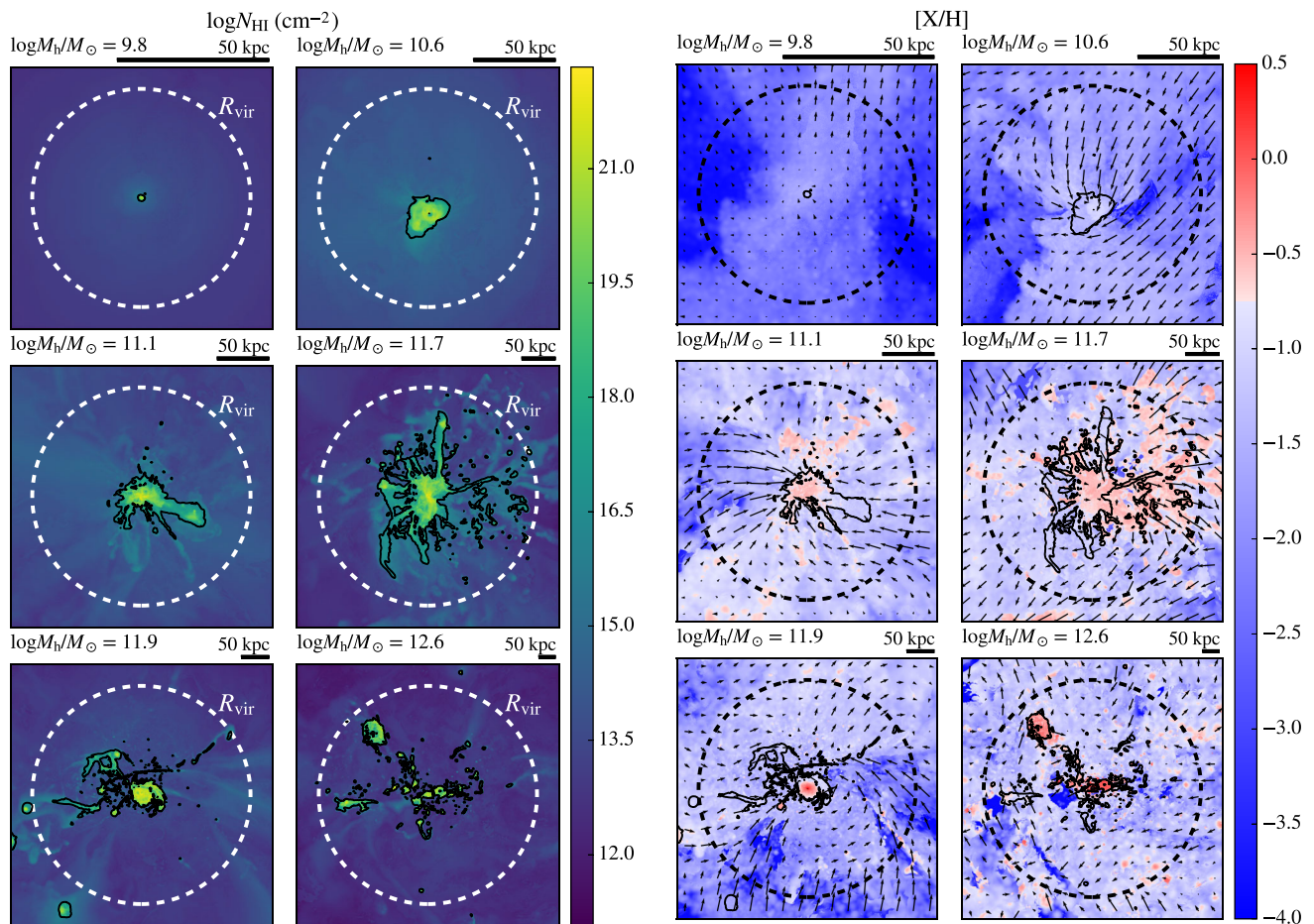
## 2.2 Extraction of absorber statistics

For each simulation, we centre our analysis on the most massive main halo within the zoom-in region, and include 20 snapshots evenly spaced in redshift in the range  $0 \leq z < 0.5$  and 11 evenly spaced snapshots in the range  $0.5 \leq z \leq 1$ . We use Amiga Halo Finder (AHF; Knollmann & Knebe 2009) and the virial overdensity definition from Bryan & Norman (1998) to identify haloes. We denote the virial radius  $R_{\text{vir}}$ . Fig. 1 shows H I and metallicity maps for six of our simulated haloes at  $z=0.5$ . From the panels on the right-hand side, which show velocity vectors overlaid on top of the metallicity maps, we anticipate a complex relationship between instantaneous velocity and metallicity in the CGM of our simulated haloes at this redshift, a point to which we will return to quantitatively in Section 4.

To extract the properties of individual CGM absorbers, we first grid the SPH particle data on to cubic 3D meshes, each with  $512^3$  cells. The full SPH kernel is used to interpolate particle data on to the grid. In cases where the grid does not resolve individual particles, the cloud-in-cell method is used to distribute the particle mass to the nearest eight grid points in a mass-conserving manner. We verified the convergence of our results with respect to grid resolution (see Appendix A). At  $z \leq 1$ , our simulations indicate that the vast majority of LLSs occur within the virial radius of haloes. To capture this halo cross-section, we use grids of side length  $2.4 R_{\text{vir}}$  for each main halo. In any given snapshot, we find that up to  $\sim 20$  per cent of LLSs in our grids are found  $> R_{\text{vir}}$  away from the central galaxy, but almost all are associated with nearby galaxies that have their own main haloes. Thus, such LLSs are accounted for by our convolutions over the dark matter halo mass function described in Section 3. The H I density for each cell is computed using a self-shielding-corrected photoionization rate (Rahmati et al. 2013), which is a function of the total hydrogen density and the UV background. We do this by post-processing our simulated grids, but have verified that computing the H I density on the particles and then depositing that density into cells gives similar results.

For each LOS, we extract the properties of the strongest H I absorber identified in velocity space. To select the strongest H I component along each LOS, we use the LOS  $\log N_{\text{H I}}$  profile binned, as a function of LOS velocity. The strongest component is identified by the peak  $N_{\text{H I}}$  along the LOS, and its width is set by the points at which  $\log N_{\text{H I}}$  drops below its mean value along the LOS. We experimented with different quantitative criteria for automatically identifying strongest absorbers along each LOS but found that there is, in general, no ambiguity in our analysis. This is because there is typically a single strongest absorber along each LOS, well separated in velocity space from other column density peaks. This approach is a good proxy for how LLSs are identified in





**Figure 1.** Left-hand panel: H I column density maps at  $z = 0.5$  for simulations m10, m10h573, m11h383, m11.9a, m12i and m13 in reading order from the top left- to bottom right-hand side. Black contours indicate  $N_{\text{HI}} \geq 10^{16.2} \text{ cm}^{-2}$  (partial LLSs and higher). Right-hand panel: Same as left-hand panel, but metallicity in solar units for the same haloes as on the left-hand side. The arrows indicate projected velocity vectors in units of the halo circular velocity. The H I column density and metallicity shown in these maps are for the strongest absorber along each sight line, calculated as described in Section 2.2. Overall, the circum-galactic gas dynamics and their relationship to gas metallicity are complex.

observations (e.g. Lehner et al. 2013) and is more accurate than, for example, estimating absorber metallicities by projecting the total gas and metal masses across the simulation volume. The latter approach can be significantly biased when the metallicity is not uniform along the LOS, which is often the case in the multiphase CGM. We do not analyse the properties of weaker absorbers. Fig. 2 illustrates this procedure for a random LOS through the m11.9a halo at  $z = 0.5$ .

For each absorber identified in this way, we evaluate the H I column density  $N_{\text{HI}}$ , metallicity  $[X/H]$ , and the gas mass-weighted radial velocity  $v_r$  of the component relative to the central galaxy. We define the metallicity  $[X/H]$  of an absorber as  $\log(Z/Z_{\odot})$ , where  $Z$  is the metal mass fraction of the absorber gas and  $Z_{\odot} = 0.014$  is the solar metal mass fraction (Asplund et al. 2010).

### 3 COSMOLOGICAL H I STATISTICS

Our zoom-in simulations provide us with predictions for the distribution of absorbers within the virial radius of galaxies as a function of halo mass and redshift. Observations that randomly select quasars on the sky, on the other hand, probe a cosmological distribution of absorbers that receives contributions from haloes of all masses. In

this section, we convolve our zoom-in results with the dark matter halo mass function to derive predictions for the cosmological incidence and column density distribution of H I CGM absorbers.

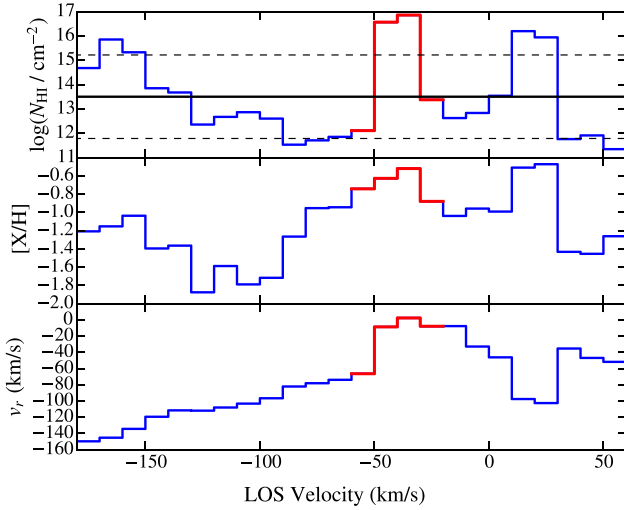
#### 3.1 Weighting haloes as a function of mass and redshift

Consider a prescribed  $N_{\text{HI}}$  range. We define  $dN_{\text{abs}}/dz$  to be the mean number of absorbers in this range per sight line per unit redshift. For a given redshift  $z$ , we quantify the contribution of haloes of different mass bins to the total  $dN_{\text{abs}}/dz$  using the weights

$$w_{\text{abs}}(z, M_{\text{rmh}}) \equiv \frac{d^2 N_{\text{abs}}}{dz d \log M_{\text{h}}} \quad (1)$$

We evaluate  $w_{\text{abs}}(z, M_{\text{h}})$  as follows.

For any main halo, we define the halo cross-section as  $\pi R_{\text{vir}}^2$ . The standard dark matter halo mass function provides the mean number density of haloes per halo mass bin as a function of redshift,  $dn_{\text{h}}/d \log M_{\text{h}}$ . The 3D halo number density can be converted into a surface density per unit redshift using the cosmological line element,  $d/dz$ . We can then simply evaluate the mean number of haloes



**Figure 2.** Example sight line profiles, as a function of LOS velocity, used to identify and quantify the properties of CGM absorbers. This is a random sight line through the main halo of our m11.9a simulation, with zero LOS velocity corresponding to the galaxy velocity. Top panel:  $\log$  of H I column density. The horizontal solid black line indicates the average value of  $\log N_{\text{HI}}$  over the sight line, and the dashed lines show  $\pm 1\sigma$  intervals. Middle panel: Metallicity profile in solar units. Bottom panel: Gas mass-weighted radial velocity relative to the central galaxy. For each sight line, we select the strongest H I absorber (highlighted in red) using the shape of the  $\log N_{\text{HI}}$  profile and extract the corresponding metallicity and radial velocity.

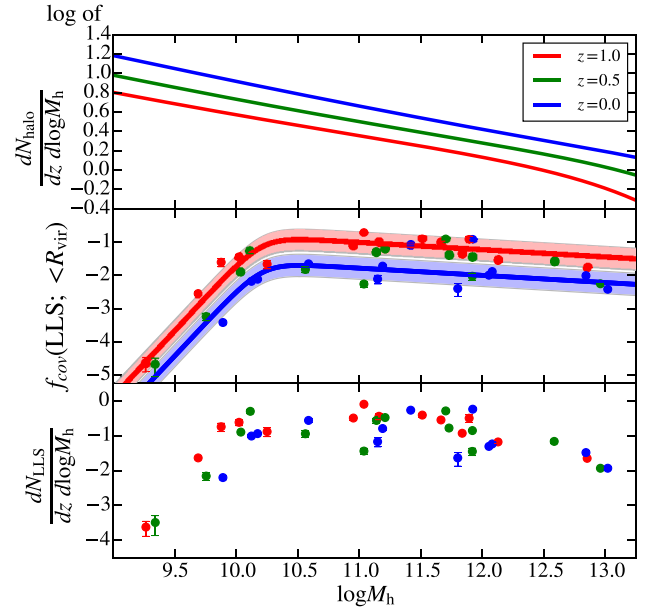
intersected per unit redshift per halo mass bin for random sight lines:

$$\frac{d^2 N_h}{dz d \log M_h} = \pi R_{\text{vir}}^2 \frac{dl}{dz} \frac{dn_h}{d \log M_h}. \quad (2)$$

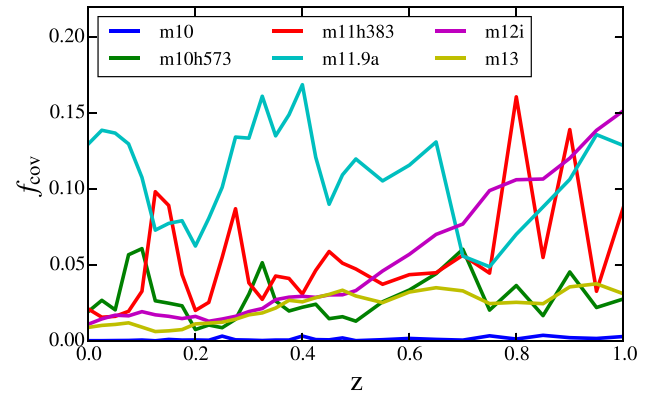
The top panel of Fig. 3 shows this quantity at  $z = 0, 0.5$  and 1. For the halo mass function, we used the CPMSO+AHF fit to  $N$ -body cosmological simulations from Watson et al. (2013). As expected, the curve increases with decreasing redshift (as haloes grow) and decreases with increasing halo mass (as haloes become rarer).

To predict the contribution of different haloes to the cosmological incidence of absorbers, we need to know the fraction of the halo cross-section  $\pi R_{\text{vir}}^2$  covered by the absorbers of interest. We define this quantity as the covering fraction  $f_{\text{cov, abs}}(< R_{\text{vir}})$  and use our zoom-in simulations to predict its values. For each simulation snapshot, we evaluate the covering fraction by extracting absorbers along one LOS for each grid pixel with impact parameter  $< R_{\text{vir}}$ . To capture viewing angle variance, we compute covering fractions for three orthogonal sky projections for each snapshot. For each halo, we plot in the middle panel of Fig. 3 the average covering fraction and the estimated standard deviation among the projections using an error bar. For this figure, we include all systems in the range  $16.2 \leq \log N_{\text{HI}} \leq 19$  to match the observational analysis of Lehner et al. (2013). This panel shows that the LLS covering fraction versus halo mass at  $z \leq 1$  peaks broadly for  $10^{10} \lesssim M_h \lesssim 10^{12} M_\odot$ . A similar (but quantitatively different) trend for LLS covering fractions peaking in relatively massive haloes is seen in the FIRE simulations at  $z = 2-4$  (Faucher-Giguère et al. 2015; Faucher-Giguère et al. 2016).

In all snapshots, the variance in  $f_{\text{cov}}$  (the error bars in Fig. 3) owing to different viewing angles is small compared to the variance arising from time variability and halo-to-halo scatter. To demonstrate the strong time variability of the covering fractions in individual



**Figure 3.** Top panel: Incidence of dark matter haloes (virial cross-sections per sight line) per unit redshift and halo mass bin. Middle panel: Covering fraction of LLSs ( $16.2 \leq \log N_{\text{HI}} \leq 19$ ) within a projected virial radius as a function of halo mass, as predicted by our zoom-in simulations. A redshift-dependent fit to the covering fractions is shown at  $z = 0$  and 1 in solid blue and red, respectively, with the shaded regions showing the standard deviation of the fit. Bottom panel: Incidence of LLS per unit redshift per unit  $\log M_h$ . The majority of randomly selected LLS are associated with haloes in the mass range  $10^{10} \lesssim M_h \lesssim 10^{12} M_\odot$ . In each panel, the solid dots correspond to the covering fractions averaged over three orthogonal sky projections and the error bars (usually smaller than the solid dot) show the sample standard deviation among the different projections.



**Figure 4.** Covering fractions for systems in the range  $16.2 \leq \log N_{\text{HI}} \leq 19$  as a function of redshift for the six simulated haloes shown in Fig. 1. The burstiness of star formation and galactic winds in our simulations drives the time variability of the covering fractions.

haloes, we plot in Fig. 4  $f_{\text{cov}}$  as a function of redshift for the six haloes shown in Fig. 1. The curves in Fig. 4 are based on the 31 simulation snapshots between  $z = 1$  and 0 analysed in this paper for each simulation. The time dependence of the covering fractions arises from the time dependence of cosmological inflows inherited from the evolution of large-scale cosmic structure and from the bursty nature of star formation and galactic winds predicted by our simulations (e.g. Faucher-Giguère et al. 2015; Muratov et al. 2015;

Sparre et al. 2015). In addition to the time variability in individual haloes, there could be some differences in time-averaged covering fractions between different haloes, for example, connected to the environments of haloes. In Section 4.3, we perform a bootstrapping analysis to quantify the uncertainty in the metallicity distribution predicted using our set of zoom-in simulations arising from halo-to-halo scatter.

We fit our simulated covering fraction data to a function of halo mass with a redshift-dependent double power law of the form

$$f_{\text{cov}}(\text{LLS}; < R_{\text{vir}}) = \frac{2f_{\text{cov},*}(1+z)^{\alpha}}{(M_{\text{h}}/M_{*})^{-\beta} + (M_{\text{h}}/M_{*})^{\gamma}}. \quad (3)$$

Performing a least-squares fit to all the  $f_{\text{cov}}$  data in the redshift interval  $0 \leq z \leq 1$  (for  $16.2 \leq \log N_{\text{HI}} \leq 19$ ) results in the best-fitting parameters  $(f_{\text{cov},*}, \log M_{*}, \beta, \gamma, \alpha) = (0.012 \pm 0.001, 10.22 \pm 0.02, 3.9 \pm 0.1, 0.22 \pm 0.03, 2.5 \pm 0.2)$ . The root-mean-square (rms)  $\log f_{\text{cov}}$  error for the fit is 0.34. The fit (plus and minimum the rms error) is plotted in the middle panel of Fig. 3 for  $z = 0$  and 1. To match the LLS definition used in previous high-redshift FIRE papers (Faucher-Giguère et al. 2015; Faucher-Giguère et al. 2016), we also fitted the covering fractions for  $\log N_{\text{HI}} \geq 17.2$ . The shape and redshift evolution of the best fit are consistent with the one above, so we held the  $\log M_{*}, \beta, \gamma$  and  $\alpha$  parameters fixed and solved for the best-fitting normalization  $f_{\text{cov},*}$  for the different  $N_{\text{HI}}$  cut. The best-fitting normalization for  $\log N_{\text{HI}} \geq 17.2$  is  $f_{\text{cov},*} = 0.0101 \pm 0.0003$ , with an rms error of  $\log f_{\text{cov}}$  of 0.30. This is 20 per cent lower than for  $16.2 \leq \log N_{\text{HI}} \leq 19$  because higher column systems are generally rarer.

The contribution of different haloes to the cosmological incidence of absorbers can then be expressed in terms of functions describing the covering fraction as a function of redshift and halo mass and the halo mass function:

$$w_{\text{abs}}(z, M_{\text{h}}) = f_{\text{cov,abs}}(< R_{\text{vir}}; z, M_{\text{h}}) \frac{d^2 N_{\text{h}}}{dz d \log M_{\text{h}}}. \quad (4)$$

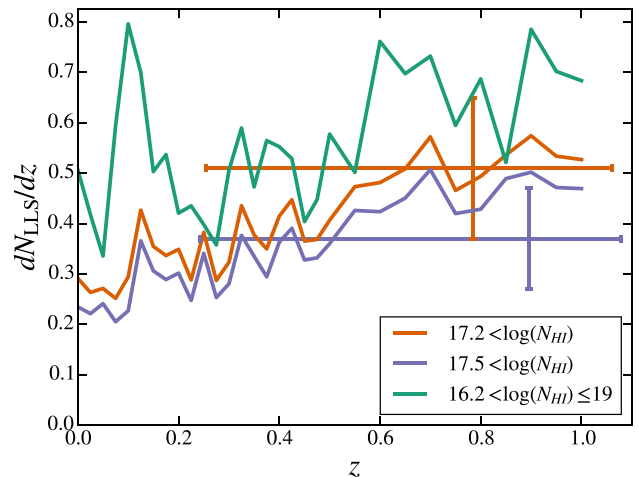
For LLSs,  $w_{\text{abs}}$  is equivalent to  $d^2 N_{\text{LLS}}/dz d \log M_{\text{h}}$ . We show  $d^2 N_{\text{LLS}}/dz d \log M_{\text{h}}$  predicted by our simulations at  $z = 0, 0.5$  and 1 in the bottom panel of Fig. 3. The bottom panel shows that, at  $z \leq 1$ , the majority of randomly selected LLSs arise from haloes in the mass range  $10^{10} \lesssim M_{\text{h}} \lesssim 10^{12} M_{\odot}$ , where covering fractions peak.

### 3.2 Cosmological incidence

The total cosmological incidence of absorbers is obtained by integrating  $w_{\text{abs}}$  over halo mass:

$$l(z) \equiv \frac{dN_{\text{abs}}}{dz} = \int d \log M_{\text{h}} w_{\text{abs}}(z, M_{\text{h}}). \quad (5)$$

We show the cosmological incidence of absorbers as a function of redshift for three different  $N_{\text{HI}}$  ranges in Fig. 5. The first range,  $16.2 \leq \log N_{\text{HI}} \leq 19$ , is chosen to match the selection from Lehner et al. (2013). The two other ranges,  $\log N_{\text{HI}} \geq 17.2$  and  $\log N_{\text{HI}} \geq 17.5$ , match the observational measurements reported in Ribaudo et al. (2011a). The comparison with the observational compilation of Ribaudo et al. (2011a) shows that convolving the covering fractions predicted by our zoom-in simulations with the halo mass function yields cosmological incidences of LLSs that are consistent with observations at  $z \leq 1$ . Overall, the simulated  $l(z)$  slightly increases from  $z = 0$  to 1; the significant redshift fluctuations apparent in Fig. 5 reflect the strong time variability of covering fractions in individual simulations in our zoom-in sample (Fig. 4).



**Figure 5.** Solid curves: The  $z \leq 1$  cosmological incidence,  $l(z)$ , of LLSs predicted by our simulations for different  $\text{H I}$  column ranges ( $10^{16.2} \leq N_{\text{HI}} \leq 10^{19}$  in green,  $N_{\text{HI}} \geq 10^{17.2}$  in orange and  $N_{\text{HI}} \geq 10^{17.5}$  in blue). The error bars show observational measurements of  $l(z)$  from the LLS census of Ribaudo et al. (2011a).

### 3.3 Column density distribution

Our zoom-in simulations can also be used to predict the column density distribution function of CGM absorbers,  $f(N_{\text{HI}})$ , defined as the incidence of absorption systems per unit  $N_{\text{HI}}$  per unit absorption length, i.e.

$$f(N_{\text{HI}}) \equiv \frac{d^2 N_{\text{abs}}}{dN_{\text{HI}} dX}. \quad (6)$$

The absorption length,  $X$ , is defined such that the incidence of absorbers per absorption length,  $l(X)$ , is constant if the comoving surface area of the absorbers is constant across cosmic time (Bahcall & Peebles 1969):

$$l(X) dX = l(z) dz, \quad (7)$$

where

$$dX = dz (1+z)^2 H_0 / H(z). \quad (8)$$

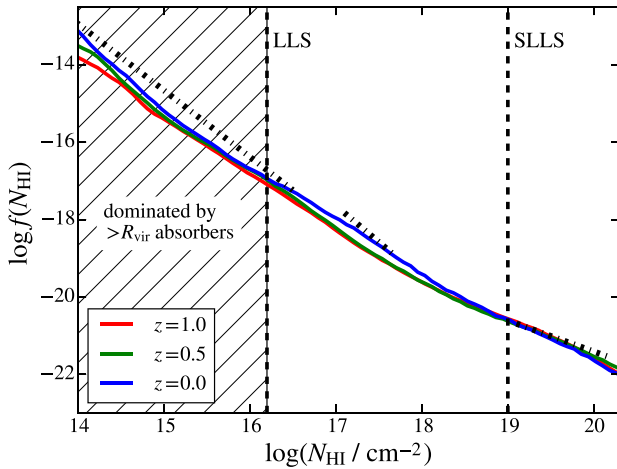
Using the procedure of the previous section to compute  $l(z)$  for small column density intervals, we can compute  $f(N_{\text{HI}})$ :

$$f(N_{\text{HI}}) \approx \frac{\Delta l(z)}{\Delta N_{\text{HI}}} \frac{dz}{dX}, \quad (9)$$

where  $\Delta l(z)$  is the contribution to the incidence of absorbers for systems with column between  $N_{\text{HI}}$  and  $N_{\text{HI}} + \Delta N_{\text{HI}}$  at redshift  $z$ .

Fig. 6 shows the column density distribution predicted by our simulations for  $10^{14} \leq N_{\text{HI}} \leq 2 \times 10^{20} \text{ cm}^{-2}$ . These predictions are compared to the observational constraints on the column density distribution from Ribaudo et al. (2011a). Overall, the predicted column density distribution agrees well with observations for LLSs and SLLSs but the simulations underpredict the column density distribution of lower column systems by  $\gtrsim 0.5$  dex at  $N_{\text{HI}} \sim 10^{15} \text{ cm}^{-2}$ . Such a discrepancy at lower columns is expected since our analysis considers only absorbers that arise within the projected virial radius of main haloes, i.e. it ignores absorbers that arise in the IGM outside dark matter haloes.





**Figure 6.** H I column density distribution,  $f(N_{\text{HI}})$ , predicted using our simulations at  $z = 0, 0.5$  and  $1$ . The black dash-dotted lines show observational constraints from Ribaud et al. (2011a). In the LLS and SLLS regimes, our simulations agree well with the observational constraints. For lower column systems, our simulations lie below the observational constraints by up to  $\gtrsim 0.5$  dex but this is because many such low-column absorbers arise outside the virial radius of galaxy haloes and therefore are not accurately captured by our analysis, which focuses on  $<R_{\text{vir}}$  absorbers.

#### 4 RELATIONSHIPS BETWEEN H I COLUMN, METALLICITY AND RADIAL KINEMATICS

We now quantify the relationships between the H I column density of absorbers, their metallicity and their radial kinematics relative to central galaxies.

For this analysis, we define ‘cosmological histograms’ that enable us to study correlations between absorbers properties properly weighted by their cosmological incidence. For any prescribed  $N_{\text{HI}}$  interval, consider a random absorber (along the sight line to a random background quasar) with properties (e.g. H I column, metallicity and radial velocity) described by the vector  $\mathbf{Y}$ . Following the formalism of Section 3, the probability density function  $P(\mathbf{Y})$  is

$$P(\mathbf{Y}) = \frac{\int \int dz d\log M_h P(\mathbf{Y}|z, \log M_h) w_{\text{abs}}}{\int \int dz d\log M_h w_{\text{abs}}}. \quad (10)$$

Here,  $P(\mathbf{Y}|z, \log M_h)$  is the conditional probability density for  $\mathbf{Y}$  given a redshift  $z$  and halo mass  $M_h$ . In our case,  $P(\mathbf{Y}|z, \log M_h)$  is the joint probability density function of H I column density, metallicity and radial velocity. We use three orthogonal sky projections for each snapshot in evaluating  $P(\mathbf{Y}|z, \log M_h)$ , although the distribution varies relatively little for different viewing angles. Since  $\mathbf{Y}$  is, in general, multivariate, we can consider different 2D projections to quantify correlations between absorber attributes.

As a complement to the 2D cosmological histograms, we compute corresponding 2D histograms that quantify the mean halo mass contributing to different bins in parameter space:

$$\langle \log M \rangle = \frac{\int \int dz d\log M_h \log M_h P(\mathbf{Y}|z, \log M_h) w_{\text{abs}}}{\int \int dz d\log M_h P(\mathbf{Y}|z, \log M_h) w_{\text{abs}}}. \quad (11)$$

For each bin of  $\mathbf{Y}$ , this histogram gives the mean  $\log M_h$  of haloes contributing to the bin weighted by the cosmological incidence of absorbers.

For any given column density interval (such as LLSs), the weighting by  $w_{\text{abs}}$  in equation (10) is derived such that the resulting 2D histogram represents the bivariate distribution of physical properties for absorbers in that column density interval along random sight

lines in the Universe. Below, we are also interested in the metallicity distribution as a function of H I column. For this, we replace the weighting factor  $w_{\text{abs}}$  by simply  $d^2 N_h / dz d \log M_h$  in equations (10) and (11). This is because in this case  $P(\mathbf{Y})$  is not restricted to a prescribed  $N_{\text{HI}}$  interval, so we need not include the  $f_{\text{cov,abs}}$  term in  $w_{\text{abs}}$ .

All histograms are evaluated using all  $0 \leq z \leq 1$  simulated data. Quantities are interpolated linearly with respect to  $z$  and  $\log M_h$  to fill gaps between snapshots. Given the time variability of the CGM in individual haloes (e.g. Fig. 4) and our modest-sized sample of simulated haloes, averaging over this full redshift interval is necessary to mitigate stochastic fluctuations. Furthermore, LLS measurements at  $z \leq 1$  are presently limited to a few dozen systems so their distributions of physical parameters are obtained by averaging over a similar redshift interval (e.g. Lehner et al. 2013; Wotta et al. 2016).

#### 4.1 Metallicity versus H I column

Fig. 7 shows the cosmological histogram for metallicity versus H I column and the corresponding mean halo mass histogram. As the panel on the left reveals, simulated LLSs typically have a metallicity within 1 dex of the median metallicity ( $[X/H] \approx -0.9$ ) of all LLSs. On the other hand, lower column systems,  $N_{\text{HI}} \lesssim 10^{15} \text{ cm}^{-2}$  have a metallicity distribution that extends well below  $[X/H] = -2$ , consistent with many of these absorbers representing a tail of IGM gas that has not yet been significantly enriched by galaxies. In the next section, we show that most  $[X/H] < -2$  absorbers have infall velocities and thus predominantly trace accretion of metal-poor intergalactic gas.

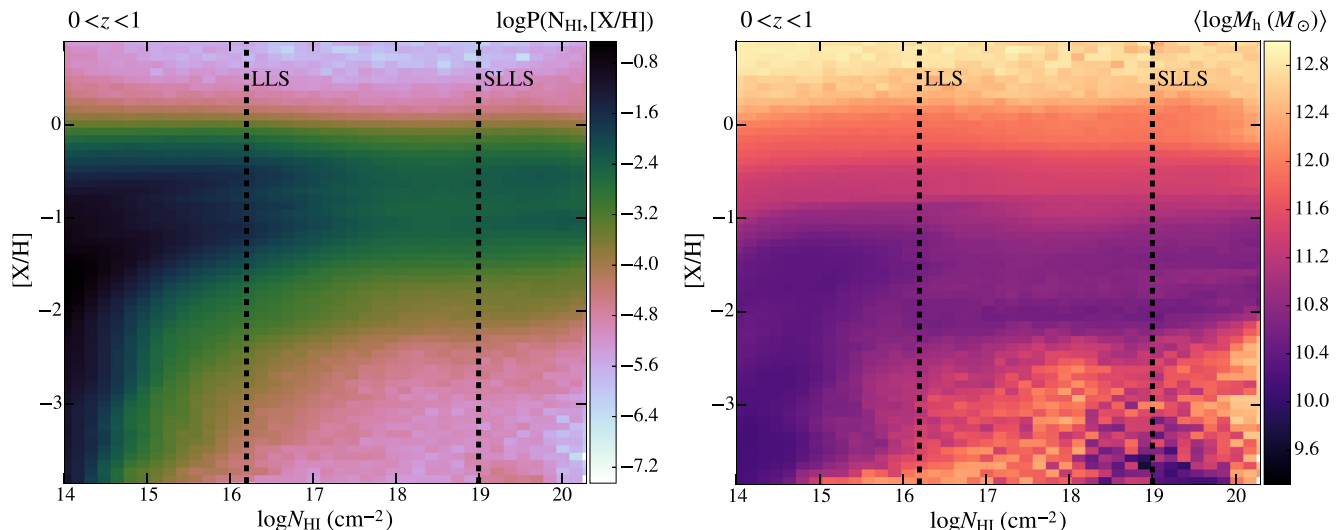
Overall, the mean mass histogram on the right-hand side of Fig. 7 reveals that more massive haloes have a contribution weighted towards higher absorber metallicities. This overall trend is a CGM analogue of the well-known mass–metallicity relation for galaxies (e.g. Tremonti et al. 2004; Zahid, Kewley & Bresolin 2011; Peeples et al. 2014) and which the FIRE simulations broadly reproduce (Ma et al. 2016). The  $[X/H] < -2$  tail of the metallicity distribution is weighted towards  $M_h \lesssim 10^{11} M_{\odot}$  haloes.

#### 4.2 Metallicity versus radial kinematics

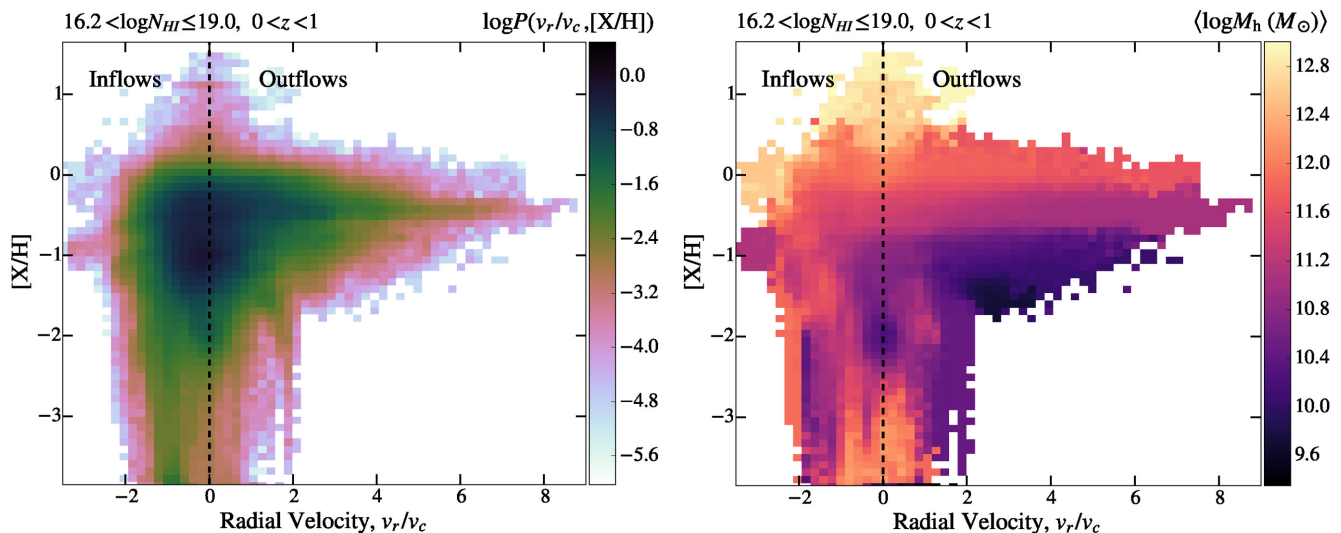
A primary motivation for our analysis is to understand how LLS metallicity can be used as a diagnostic of inflows and outflows. Fig. 8 illustrates the relationship between absorber metallicity and its radial velocity relative to the central galaxy,  $v_r$ , along with the corresponding mean mass histogram, for absorbers in the range  $16.2 \leq \log N_{\text{HI}} \leq 19$  (i.e. including partial LLSs). Positive radial velocities correspond to gas that is outflowing in an instantaneous sense, while negative radial velocities correspond to gas that is inflowing. In order to compare radial velocities for absorbers arising in a wide range of halo mass, velocities are normalized by halo circular velocity,  $v_c = \sqrt{GM_h/R_{\text{vir}}}$ , which corresponds to  $\sim (30, 130, 270) \text{ km s}^{-1}$  for  $M_h \sim (10^{10}, 10^{12}, 10^{13}) M_{\odot}$ .

Fig. 8 shows that for cosmologically selected LLSs in the redshift interval  $0 \leq z \leq 1$  in our simulations:

- (i) very high velocity ( $v_r/v_c > 2$ ) outflows tend to have high-metallicity  $[X/H] \approx -0.5$ ;
  - (ii) very low metallicity ( $[X/H] < -2$ ) LLSs tend to have radial velocities corresponding to inflows ( $v_r < 0$ );
  - (iii) most LLSs have modest absolute velocity  $|v_r/v_c| \leq 2$ .
- Overall, these LLSs have a broad metallicity distribution ranging



**Figure 7.** Left-hand panel: Cosmological 2D histogram for metallicity versus H I column density from our simulations (representative of randomly selected LLSs). Simulated LLSs typically have a metallicity within 1 dex of the LLS mean metallicity  $[X/H] \approx -0.9$ . The metallicity distribution for lower column systems,  $N_{\text{HI}} \lesssim 10^{15} \text{ cm}^{-2}$ , extends to lower metallicities. Right-hand panel: 2D histogram showing the mean  $\log M_h$  contributing to different bins on the left-hand side. Overall, the contribution of more massive haloes is weighted towards higher metallicities. The colour bars are logarithmically scaled. We include systems with  $N_{\text{HI}} < 10^{16.2} \text{ cm}^{-2}$  to provide information about their distribution integrated over galaxy haloes, but caution that this distribution does not include a contribution from absorbers arising outside the virial radius of haloes.



**Figure 8.** Left-hand panel: Cosmological 2D histogram for LLS metallicity versus radial velocity relative to the central galaxy (representative of randomly selected LLSs). High velocity ( $v_r/v_c > 2$ ) outflows tend to have high metallicities  $[X/H] \sim -0.5$  and very low metallicity ( $[X/H] < -2$ ) LLSs tend to have infall radial velocities. However, most LLSs occupy an intermediate central region in metallicity-radial velocity space. For these more typical LLSs, there is no clear trend between metallicity and radial kinematics. Metal-enriched inflows arise in the FIRE simulations as a result of galactic winds that efficiently recycle at low redshift. Right-hand panel: Mean mass histogram showing the mean  $\log M_h$  contributing to different bins on the left-hand side. The colour bars are logarithmically scaled.

from  $[X/H] \approx -2$  to  $[X/H] \approx 0$ , with no clear trend with radial kinematics.

Thus, our simulations suggest that very low metallicity LLSs with  $[X/H] \lesssim -2$  generally trace infalling IGM gas (consistent with observational interpretations of such low-metallicity absorbers, e.g. Fumagalli et al. 2011; Ribaud et al. 2011b) but that metallicity alone cannot robustly distinguish between inflows and outflows for the majority of LLSs with higher metallicity.

The concentration of LLSs with  $|v_r/v_c| \leq 2$  is easy to understand since  $v_c$  is both the characteristic velocity of gas that is accelerated as it falls into haloes from the IGM and the characteristic velocity of galactic winds in the FIRE simulations (Muratov et al. 2015). The difficulty of associating metal-enriched LLSs with inflows versus outflows is due in large part to the importance of wind recycling in our simulations (for a previous analysis of how wind recycling shapes the galaxy stellar mass function, see Oppenheimer et al. 2010). When galactic winds recycle efficiently, metal-rich wind ejecta that initially have positive radial velocity later fall back on to



the source galaxy with  $v_r < 0$ . Muratov et al. (2015) showed that the powerful galactic winds driven by high-redshift galaxies transform into galactic fountains by  $z \sim 0$  (or earlier) in massive galaxies in the FIRE simulations. Ma et al. (2016) further demonstrated that the FIRE galaxies, except dwarfs with stellar mass  $M_* \lesssim 10^7 M_\odot$ , retain most of the metals they have produced in their haloes. Anglés-Alcázar et al. (2016) confirm the importance of wind recycling in the FIRE simulations using a particle tracking analysis that traces the full history of baryons.

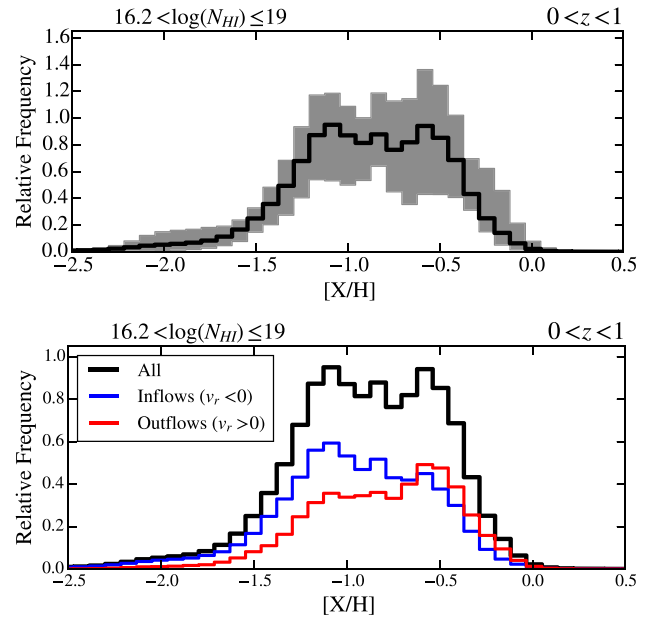
As a check of our above result based on instantaneous radial kinematics that absorbers with  $[X/H] \lesssim -2$  can generally be associated with the accretion of fresh gas from the IGM, we have used the particle tracking pipeline from Anglés-Alcázar et al. (2016) and found the following: In the main halo of our m12i simulation at  $z = 0.5$  nearly all the gas associated with LLSs and classified as fresh accretion based on its full time history has metallicity  $[X/H] \lesssim -1$ , with more than half having metallicity  $[X/H] \lesssim -2$ . On the other hand, nearly all the wind LLSs in this halo have metallicity  $[X/H] \gtrsim -2$ . We plan to more systematically extend this particle tracking analysis in future work. We note that previous authors found using particle tracking that wind recycling can constitute a significant (and possibly dominant) fraction of the instantaneously infalling CGM gas at late times in other simulations as well (Ford et al. 2014; Christensen et al. 2016). However, the magnitude of this effect depends on the properties of the (uncertain) wind models so previous results do not necessarily quantitatively apply to the FIRE simulations.

The ambiguities in using LLS metallicity to distinguish between inflows and outflows in the FIRE simulations at low redshift are in contrast to some earlier simulation analyses at high redshift ( $z \sim 3$ ), in which maps of the CGM appear to show a much stronger correlation between instantaneous radial kinematics and metallicity than the  $z = 0.5$  haloes shown in Fig. 1 or than quantified in this section (e.g. Shen et al. 2012). Quantitatively, the analysis of a single main halo in the ErisMC simulation by Shen et al. (2012) revealed a factor of  $\sim 10$  difference in mean metallicity between instantaneously inflowing and outflowing halo gas at  $z \geq 3$ . The difference between our results and the high-redshift results of Shen et al. (2012) is likely at least partially attributable to the fact that both cosmological inflows and galactic winds are much stronger at high redshift. Furthermore, high-redshift accretion of cool intergalactic gas dense enough to give rise to LLS absorption tends to be collimated in streams (e.g. Kereš et al. 2005; Dekel et al. 2009; Faucher-Giguère & Kereš 2011; Fumagalli et al. 2011) between which outflowing gas can propagate relatively unimpeded. By  $z \sim 0.5$ , Fig. 1 shows that well-defined cool streams of infalling gas have largely disappeared in the simulations. We therefore caution that our present conclusions may not apply at  $z > 1$ .

### 4.3 The metallicity distribution of LLSs

To directly compare with the LLS metallicity distribution recently measured by Lehner et al. (2013), we project the 2D histogram in Fig. 8 on to the metallicity axis. The result is shown in Fig. 9. The simulated metallicity distribution has a mean (standard deviation)  $\langle [X/H] \rangle = -0.9 (0.4)$ .

To test the statistical significance of features in the metallicity distribution, we performed a bootstrapping analysis. We recalculated the metallicity histogram 200 times, each time using a new sample of simulations drawn with replacement from our original set of simulations. The shaded region in the top panel of Fig. 9 indicates the 95 per cent confidence interval for each metallicity



**Figure 9.** Top panel: Overall metallicity distribution for  $z \leq 1$  LLSs, averaging over haloes of all masses (solid curve). The shaded region shows the 95 per cent confidence interval estimated by sampling from the full simulation set with replacement. The multiple peaks apparent in the fiducial distribution are not significant. Bottom panel: Same LLS metallicity distribution as in the top panel (black) and divided into inflow (blue) and outflow (red) components based on instantaneous radial velocity. Outflowing LLSs have on average slightly higher metallicity than inflowing ones, but the two distributions overlap strongly, with no clean separation in metallicity.

bin. From this analysis, we conclude that apparent multiple peaks in the metallicity distribution computed using our entire simulation sample are consistent with statistical noise inherent to the finite number of simulations in our sample. Thus, we do not find significant evidence for a multimodal metallicity distribution in our simulations. It is possible, though, that the simulated distribution has weak multimodal features within the noise of our analysis.

The lack of a significant metallicity bimodality in our simulations is in contrast with the possible metallicity bimodality observed by Lehner et al. (2013) and in the expanded sample of Wotta et al. (2016), which suggests a well-defined dip at  $[X/H] \approx -1$ . In our simulations, certain individual snapshots show rather well-defined bimodal metallicity distributions (for example, the m11.9a snapshot shown in Fig. 1) with the low- and high-metallicity peaks roughly associated with inflows and galactic ejecta. However, inflows and outflows are highly time variable in our simulations and such features are transients. The main result of Fig. 9 is that bimodal features in the metallicity distribution do not survive cosmological averaging over the broad range of haloes in our simulations. This result is consistent with the broad range of halo mass contributing to the cosmological incidence of LLSs. Over the halo mass range  $M_h \sim 10^{10} - 10^{12} M_\odot$  that contributes most to the cosmological LLS incidence, the interstellar gas phase and stellar metallicities span a total range of  $\sim 2$  dex in our simulations (taking into account both the systematic increase with stellar mass and scatter around the mean mass–metallicity relation). These simulations are in good agreement with observational measurements of the mass–metallicity relations (Ma et al. 2016). Since the metallicity of galactic winds is similar to the metallicity of the interstellar medium (Muratov et al. 2017), the circum-galactic LLS metallicity distribution predicted by our

simulations should therefore have a total width comparable to this. As a result, any clean bimodal ‘dip’ in the metallicity distribution at a given halo mass (if present) may be expected to be substantially washed out in the cosmological average.

Accretion of fresh gas from the IGM could, in principle, give rise to a distinct very low metallicity peak in the LLS metallicity distribution, but such a distinct component is not present in our simulated cosmological distribution. This is likely because even gas that is first accreting on to a central galaxy from the IGM tends to be pre-enriched by ejecta from surrounding lower mass galaxies. Smooth IGM accretion generally originates from similar large-scale structures as infalling satellite galaxies, and we have previously noted the effects of galactic winds from satellites on the observational properties of infalling gas at high redshift (Faucher-Giguère et al. 2015; Faucher-Giguère et al. 2016).

In the updated  $z \leq 1$  metallicity analysis of Wotta et al. (2016), the statistical evidence for a bimodality is strongest for the subsample of partial LLSs, in the range  $16.2 \leq \log N_{\text{H I}} \leq 17.2$ . However, 15 out of 44 absorbers in that sample only have lower or upper limits on their metallicity. Allowing the lower limit absorbers to have arbitrarily high metallicity and the upper limit absorbers to have arbitrarily low metallicity could significantly flatten the distribution and potentially reduce the statistical significance of the dip at  $[X/H] \approx -1$ . It will be important to measure more precisely the metallicities of absorbers with only limits currently available (or to rigorously account for the limits in the statistical analysis) to more conclusively determine statistical significance of the observed dip. At  $z > 2$ , the observational analyses of Fumagalli et al. (2016) and Lehner et al. (2016) reveal a broad unimodal LLS metallicity distribution.

Another discrepancy between our simulations and the metallicity distributions measured by Lehner et al. (2013) and Wotta et al. (2016), which cannot be explained by the presence of limits in the observational sample, is that our simulations contain significantly fewer very low metallicity LLSs than the observations. For example, our simulated metallicity distribution contains very few ( $\approx 10$  per cent) LLSs with  $[X/H] < -1.5$  but about half of Wotta et al.’s low-metallicity branch have such low metallicity. Wotta et al. explore systematic uncertainties in metallicity estimates from the assumed UV background model and find that using a UV background model with a harder spectrum due to a larger contribution from quasars would increase their inferred metallicities by up to  $\sim 0.6$  dex. This effect could explain part of the discrepancy at low metallicity if pushed to its extreme, but would introduce some tension at the high-metallicity end.

Differences in metal mixing may also contribute to the low-metallicity discrepancy. In our SPH simulations, metals are advected by particles but do not diffuse between particles, so that metal mixing occurring below the resolution limit is not captured. Underestimating subgrid metal mixing in the simulations should (if anything) overestimate the number of very low metallicity LLSs, rather than underestimate it. We have verified that the ‘missing’ low-metallicity systems cannot be explained by the absence of subgrid metal mixing in our simulations by analysing a re-simulation of our m12i halo that includes a subgrid turbulent metal diffusion model similar to Shen, Wadsley & Stinson (2010). This re-simulation did not produce more  $[X/H] < -1.5$  absorbers. On the other hand, if metals in the real CGM are poorly mixed, then cosmological simulations (both SPH and grid-based) could overestimate their mixing by forcing new metals to be distributed to at least one resolution element. In our simulations, metals produced by stellar evolution are distributed to the  $\sim 60$  gas particles in the SPH smoothing

kernel, corresponding to a gas mass  $> 10^6 M_\odot$  for the m11.4a and m11.9a simulations whose parameters are listed in Table 1. In an observational analysis of  $z \sim 2.3$  CIV absorbers, Schaye, Carswell & Kim (2007) concluded that those intergalactic absorbers likely arise from a large population of compact metal clumps with masses  $M_{\text{clump}} \sim 10^2 M_\odot$ , which we are unable to resolve in our simulations. If the metals in  $z \leq 1$  LLSs are similarly poorly mixed, then our simulations could overestimate the metallicity of many LLS sight lines within which metals in reality would not have had enough time to mix. This effect could potentially ‘hide’ a population of low-metallicity LLSs.

To further investigate the relationship between metallicity and inflows versus outflows, we divide the total metallicity distribution into instantaneously inflowing ( $v_r < 0$ ) and outflowing ( $v_r > 0$ ) components in the bottom panel of Fig. 9. The instantaneously inflowing LLSs have a mean (standard deviation)  $\langle [X/H] \rangle = -1.0$  (0.4), while the instantaneously outflowing LLSs have a mean (standard deviation)  $\langle [X/H] \rangle = -0.8$  (0.4). Consistent with our results from the previous section, infalling LLSs are on average of slightly lower metallicity than outflowing ones, but the distributions are not cleanly separated in metallicity. This suggests that metallicity alone cannot be used as a diagnostic to distinguish instantaneous inflows from instantaneous outflows for LLSs with  $[X/H] \gtrsim -2.0$ , i.e. the majority of LLSs. However, as discussed in Section 4.2, the lowest metallicity systems with  $[X/H] \lesssim -2$  do tend to be inflows.

## 5 CONCLUSIONS

We have used cosmological hydrodynamic zoom-in simulations from the FIRE project to investigate the physical properties of circum-galactic absorption at  $z \leq 1$ , with particular emphasis on LLSs and their use as diagnostics of cosmological inflows and galactic winds. The FIRE simulations self-consistently generate galactic winds from energy and momentum injection on the scale of star-forming regions by stellar feedback. We analysed 14 simulations covering the halo mass range  $M_h \sim 10^9$ – $10^{13} M_\odot$  at  $z = 0$ , which we convolved with the dark matter halo mass function to obtain cosmological statistics representative of absorbers randomly selected along unbiased sight lines to background quasars. Our main conclusions are as follows:

- (i) When convolved with the dark matter halo mass function, the FIRE simulations are consistent with the observed cosmological incidence and H I column density distribution of LLSs at  $z \leq 1$ .
- (ii) The majority of HI-selected LLSs are associated with relatively massive haloes in the mass range  $10^{10} \lesssim M_h \lesssim 10^{12} M_\odot$ . Dwarf haloes with  $M_h \lesssim 10^{9.5} M_\odot$  have extremely small ( $\lesssim 10^{-4}$ ) LLS covering fractions within a projected virial radius.
- (iii) The LLS covering fractions of individual haloes are strongly time variable. The strong variability results from a combination of time-variable inflows (including accreting satellites) and bursty outflows.
- (iv) Simulated LLSs typically have a metallicity within 1 dex of the mean metallicity  $[X/H] \approx -0.9$ . The metallicity distribution of lower column systems,  $N_{\text{H I}} \lesssim 10^{15} \text{ cm}^{-2}$ , extends well below  $[X/H] = -2$ . This is consistent with many of these low-column absorbers representing a tail of IGM gas that has not yet been significantly enriched by galaxies.
- (v) LLSs with large radial outflow velocity ( $v_r/v_c > 2$ ) tend to have high metallicities  $[X/H] \sim -0.5$ , while very low metallicity ( $[X/H] < -2$ ) LLSs tend to have infall radial velocities.

(vi) Most LLSs have moderate radial velocities ( $|v_r/v_c| \leq 2$ ). For these more common LLSs, there is no strong trend between metallicity and instantaneous radial kinematics.

(vii) When separating LLSs in groups with  $v_r < 0$  (inflows) and  $v_r > 0$  (outflows), the inflowing LLSs have a slightly lower mean metallicity ( $\langle [X/H] \rangle = -1.0$ ) than the outflowing LLSs ( $\langle [X/H] \rangle = -0.8$ ). However, both distributions have a standard deviation of  $[X/H] \approx 0.4$ , causing them to overlap strongly.

(viii) Overall, we find no significant evidence for a bimodality of LLS metallicity in our simulations. This result is in tension with observations that suggest two metallicity branches clearly separated at  $[X/H] \approx -1$ .

(ix) The simulated metallicity distribution also lacks a population of low-metallicity LLSs ( $[X/H] \lesssim -1.5$ ) that is prominent in observations, with only  $\approx 10$  per cent of all simulated LLSs having that low metallicity. The existence of such low-metallicity LLSs may indicate that metals are poorly mixed in the observed CGM on scales below the resolution of our simulations.

Overall, our simulations indicate that very low metallicity LLSs are predominantly associated with cosmological inflows at  $z \leq 1$ , but that higher metallicity LLSs trace gas with roughly equal probability of having instantaneous inflow or outflow kinematics. This result is largely a consequence of the prevalence of gas recycling in the FIRE simulations, which causes a large fraction of metal-rich galactic wind ejecta to later fall back on to galaxies (Anglés-Alcázar et al. 2016). Thus, metallicity is a powerful diagnostic of pristine intergalactic inflows but, in general, cannot robustly distinguish between recent outflows and inflows of recycled gas. It will be interesting to sharpen this result by using particle tracking to more accurately identify the physical nature and history of gas elements in our simulations. Such an analysis may reveal that metallicity is a more powerful diagnostic of gas that has been previously processed by galaxies versus not than of instantaneous inflows versus instantaneous outflows. It will also be important to firm up the statistical significance of our analysis using a larger sample of simulated haloes.

We also plan to extend our analysis of  $z \leq 1$  LLSs to the high-redshift Universe, where observational measurements of LLSs and their metallicities are also available (Lehner et al. 2014; Fumagalli et al. 2016; Lehner et al. 2016). At high redshift, many LLSs arise outside the virial radius of galaxy haloes (e.g. Cooper et al. 2015) and the present zoom-in approach will not be adequate to study cosmological statistics. We plan to pursue full-volume cosmological simulations with the FIRE resolution and physics to address such IGM questions. Another promising observational diagnostic of inflows and outflows that will be worthwhile to investigate using simulations is the distribution in azimuthal angle relative to the galactic disc plane of strong Mg II absorbers (e.g. Bordoloi et al. 2011; Bouché et al. 2012; Kacprzak, Churchill & Nielsen 2012; Lan, Ménard & Zhu 2014).

## ACKNOWLEDGEMENTS

The authors are grateful to Nicolas Lehner, Chris Wotta, Chris Howk, J. X. Prochaska, Cameron Liang, Joop Schaye and Andrey Kravtsov for discussions regarding the observed Lyman limit system metallicity bimodality. We are also grateful to Benedikt Diemer for providing a set of PYTHON cosmology modules used in this work. ZH, CAFG and DAA were supported by NSF through grants AST-1412836, AST-1517491 and DGE-0948017, by NASA through grant NNX15AB22G, and by STScI through grants

HST-AR-14293.001-A and HST-GO-14268.022-A. DK and TKC were supported in part by NSF grant AST-1412153 and Cottrell Scholar Award from the Research Corporation for Science Advancement. Support for PFH was provided by an Alfred P. Sloan Research Fellowship, NASA ATP grant NNX14AH35G, and NSF grants AST-1411920 and AST-1455342. EQ was supported by NASA ATP grant 12-ATP-120183, a Simons Investigator award from the Simons Foundation and the David and Lucile Packard Foundation. The simulations analysed in this paper were run on XSEDE computational resources (allocations TG-AST120025, TG-AST130039, and TG-AST140023), on the NASA Pleiades cluster (allocation SMD-14-5189), on the Northwestern Quest computer cluster, and on the Caltech Zwicky computer cluster.

## REFERENCES

- Agertz O. et al., 2007, *MNRAS*, 380, 963  
 Aguirre A., Hernquist L., Schaye J., Weinberg D. H., Katz N., Gardner J., 2001, *ApJ*, 560, 599  
 Altay G., Theuns T., Schaye J., Crighton N. H. M., Dalla Vecchia C., 2011, *ApJ*, 737, L37  
 Altay G., Theuns T., Schaye J., Booth C. M., Vecchia C. D., 2013, *MNRAS*, 436, 2689  
 Anglés-Alcázar D., Davé R., Özel F., Oppenheimer B. D., 2014, *ApJ*, 782, 84  
 Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Hopkins P. F., Quataert E., Murray N., 2016, *MNRAS*, preprint ([arXiv:1610.08523](https://arxiv.org/abs/1610.08523))  
 Asplund M., Grevesse N., Sauval A. J., Scott P., 2010, *Astrophys. Space Sci.*, 328, 179  
 Bahcall J. N., Peebles P. J. E., 1969, *ApJ*, 156, L7  
 Barnes J. E., 2012, *MNRAS*, 425, 1104  
 Bauermeister A., Blitz L., Ma C.-P., 2010, *ApJ*, 717, 323  
 Bird S., Vogelsberger M., Sijacki D., Zaldarriaga M., Springel V., Hernquist L., 2013, *MNRAS*, 429, 3341  
 Booth C. M., Schaye J., Delgado J. D., Dalla Vecchia C., 2012, *MNRAS*, 420, 1053  
 Bordoloi et al., 2011, *ApJ*, 743, 10  
 Bouché N., Hohensee W., Vargas R., Kacprzak G. G., Martin C. L., Cooke J., Churchill C. W., 2012, *MNRAS*, 426, 801  
 Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, *ApJ*, 694, 396  
 Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80  
 Chan T. K., Kereš D., Oñorbe J., Hopkins P. F., Muratov a. L., Faucher-Giguère C. a., Quataert E., 2015, *MNRAS*, 454, 2981  
 Cicone et al., 2014, *A&A*, 562, A21  
 Christensen C. R., Davé R., Governato F., Pontzen A., Brooks A., Munshi F., Quinn T., Wadsley J., 2016, *ApJ*, 824, 57  
 Cooper T. J., Simcoe R. A., Cooksey K. L., O'Meara J. M., Torrey P., 2015, *ApJ*, 812, 58  
 Dekel A., Birnboim Y., 2006, *MNRAS*, 368, 2  
 Dekel A. et al., 2009, *Nature*, 457, 451  
 Faucher-Giguère C. a., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416  
 Faucher-Giguère C. A., Kereš D., 2011, *MNRAS*, 412, 118  
 Faucher-Giguère C.-A., Kereš D., Ma C.-P., 2011, *MNRAS*, 417, 2982  
 Faucher-Giguère C.-A., Hopkins P. F., Kere D., Muratov A. L., Quataert E., Murray N., 2015, *MNRAS*, 449, 987  
 Faucher-Giguère C.-A., Feldmann R., Quataert E., Kereš D., Hopkins P. F., Murray N., 2016, *MNRAS*, 461, L32  
 Feldmann R., Quataert E., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2016, *MNRAS*, preprint ([arXiv:1610.02411](https://arxiv.org/abs/1610.02411))  
 Ford A. B., Davé R., Oppenheimer B. D., Katz N., Kollmeier J. A., Thompson R., Weinberg D. H., 2014, *MNRAS*, 444, 1260  
 Fox A. J. et al., 2013, *ApJ*, 778, 187  
 Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J. R., 2011, *MNRAS*, 418, 1796

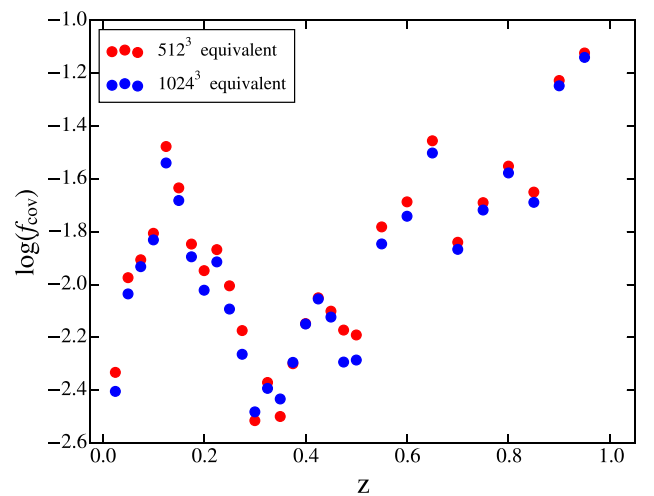


Fumagalli M., O'Meara J. M., Prochaska J. X., 2011, *Science*, 334, 1245  
 Fumagalli M., O'Meara J. M., Prochaska J. X., Worseck G., 2013, *ApJ*, 775, 78  
 Fumagalli M., Hennawi J. F., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J., 2014, *ApJ*, 780, 74  
 Fumagalli M., O'Meara J. M., Prochaska J. X., 2016, *MNRAS*, 455, 4100  
 Goerdt T., Dekel A., Sternberg A., Gnat O., Ceverino D., 2012, *MNRAS*, 424, 2292  
 Greene J. E., Zakamska N. L., Ho L. C., Barth A. J., 2011, *ApJ*, 732, 9  
 Gurvich A., Burkhart B., Bird S., 2016, *ApJ*, 875, 175  
 Heckman T. M., Lehnert M. D., Strickland D. K., Armus L., 2000, *ApJS*, 129, 493  
 Hernquist L., Katz N., Weinberg D. H., Miralda-Escudé J., 1996, *ApJ*, 457, L51  
 Hopkins P. F., 2013, *MNRAS*, 428, 2840  
 Hopkins P. F., 2015, *MNRAS*, 450, 53  
 Hopkins P. F., Keres D., Onorbe J., Faucher-Giguere C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, 445, 581  
 Jones T., Stark D. P., Ellis R. S., 2012, *ApJ*, 751, 51  
 Kacprzak G. G., Churchill C. W., Nielsen N. M., 2012, *ApJ*, 760, L7  
 Katz N., White S. D. M., 1993, *ApJ*, 412, 455  
 Katz N., Weinberg D. H., Hernquist L., 1996, *ApJS*, 105, 19  
 Kereš D., Katz N., Weinberg D. H., Davé R., 2005, *MNRAS*, 363, 2  
 Kereš D., Katz N., Fardal M., Davé R., Weinberg D. H., 2009, *MNRAS*, 395, 160  
 Knollmann S. R., Knebe A., 2009, *ApJS*, 182, 608  
 Kohler K., Gnedin N. Y., 2007, *ApJ*, 655, 685  
 Kollmeier J. A., Miralda-Escudé J., Cen R., Ostriker J. P., 2006, *ApJ*, 638, 52  
 Lan T.-W., Ménard B., Zhu G., 2014, *ApJ*, 795, 31  
 Lehner N. et al., 2013, *ApJ*, 770, 138  
 Lehner N., O'Meara J., Fox A., 2014, *ApJ*, 788, 119  
 Lehner N., O'Meara J. M., Howk J. C., Prochaska J. X., Fumagalli M., 2016, *ApJ*, 833, 283  
 Leitherer C. et al., 1999, *ApJS*, 123, 3  
 Liang C. J., Kravtsov A. V., Agertz O., 2016, *MNRAS*, 458, 1164  
 McQuinn M., Oh S. P., Faucher-Giguere C.-A., 2011, *ApJ*, 743, 82  
 Ma X., Hopkins P. F., Faucher-Giguere C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, *MNRAS*, 456, 2140  
 Martin C. L., 2005, *ApJ*, 621, 227  
 Muratov A. L., Keres D., Faucher-Giguere C.-A., Hopkins P. F., Quataert E., Murray N., 2015, *MNRAS*, 454, 2691  
 Muratov A. L. et al., 2017, *MNRAS*, 468, 4170  
 Murray N., 2011, *ApJ*, 729, 133  
 Oppenheimer B. D., Davé R., 2006, *MNRAS*, 373, 1265  
 Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, *MNRAS*, 406, 2325  
 Peebles M. S., Werk J. K., Tumlinson J., Oppenheimer B. D., Prochaska J. X., Katz N., Weinberg D. H., 2014, *ApJ*, 786, 54  
 Planck Collaboration XIII, 2015, *A&A*, 519, A13  
 Porter D. H., 1985, PhD thesis, California Univ., Berkeley  
 Price D. J., Monaghan J. J., 2007, *MNRAS*, 374, 1347  
 Prochaska J. X., Wolfe A. M., 2009, *ApJ*, 696, 1543  
 Prochaska J. X., O'Meara J. M., Fumagalli M., Bernstein R. A., Burles S. M., 2015, *ApJS*, 221, 2  
 Quiet S. et al., 2016, *MNRAS*, 458, 4074  
 Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, *MNRAS*, 430, 2427  
 Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, *MNRAS*, 452, 2034  
 Ribaldo J., Lehner N., Howk J. C., 2011a, *ApJ*, 736, 42  
 Ribaldo J., Lehner N., Howk J. C., Werk J. K., Tripp T. M., Prochaska J. X., Meiring J. D., Tumlinson J., 2011b, *ApJ*, 743, 207  
 Rubin K. H. R., Prochaska J. X., Koo D. C., Phillips A. C., Martin C. L., Winstrom L. O., 2014, *ApJ*, 794, 156  
 Schaye J., Carswell R. F., Kim T.-S., 2007, *MNRAS*, 379, 1169  
 Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, *ApJ*, 588, 65

Shen S., Wadsley J., Stinson G., 2010, *MNRAS*, 407, 1581  
 Shen S., Madau P., Aguirre A., Guedes J., Mayer L., Wadsley J., 2012, *ApJ*, 760, 50  
 Somerville R. S., Davé R., 2015, *Annu. Rev. Astron. Astrophys.*, 53, 51  
 Sparre M., Hayward C. C., Feldmann R., Faucher-Giguere C.-A., Muratov A. L., Kereš D., Hopkins P. F., 2015, *MNRAS*, 446, 88  
 Springel V., 2005, *MNRAS*, 364, 1105  
 Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, *ApJ*, 717, 289  
 Tremonti C. A. et al., 2004, *ApJ*, 613, 898  
 Tripp T. M. et al., 2011, *Science*, 334, 952  
 van de Voort F., Schaye J., Booth C. M., Haas M. R., Dalla Vecchia C., 2011, *MNRAS*, 414, 2458  
 van de Voort F., Schaye J., Altay G., Theuns T., 2012, *MNRAS*, 421, 2809  
 Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, *MNRAS*, 433, 1230  
 Weiner B. J. et al., 2009, *ApJ*, 692, 187  
 Wotta C. B., Lehner N., Howk J. C., O'Meara J. M., Prochaska J. X., 2016, *ApJ*, 831, 95  
 Zahid H. J., Kewley L. J., Bresolin F., 2011, *ApJ*, 730, 137

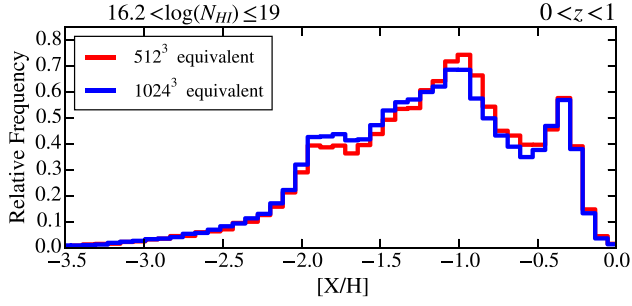
## APPENDIX A: CONVERGENCE PROPERTIES

To test our results for convergence with respect to grid resolution, we regridded a single quadrant of m12i at both fiducial resolution and at one higher resolution level over the full redshift range  $0 \leq z \leq 1$ . If we define the origin as the centre of the m12i galaxy, our new cubic meshes have  $256^3$  and  $512^3$  grid points over a volume with  $(x_{\min}, y_{\min}, z_{\min}) = (0, 0, 0)$  and  $(x_{\max}, y_{\max}, z_{\max}) = (1.2R_{\text{vir}}, 1.2R_{\text{vir}}, 1.2R_{\text{vir}})$ . This produces grids with  $512^3$  and  $1024^3$  equivalent resolution over the full halo. The results of our grid resolution convergence tests are shown in Fig. A1 for LLS covering fractions and in Fig. A2 for the metallicity distribution. Both the covering fractions and the metallicity distribution are well converged with respect to grid resolution. This is expected since the SPH smoothing lengths of LLS gas are comparable to the grid cell size at the fiducial grid resolution used for our analysis. We note that the LLS covering fractions plotted in Fig. A1 for the m12i



**Figure A1.** Grid resolution convergence for the LLS ( $16.2 \leq \log N_{\text{HI}} \leq 19$ ) covering fractions within a projected virial radius in the m12i simulation. We tested grids equivalent to our fiducial resolution and one higher resolution level ( $512^3$  and  $1024^3$  grid points on grids with side lengths of  $2.4R_{\text{vir}}$ ), but spanning a single quadrant of the virial radius. The covering fractions are well converged for the fiducial  $512^3$  grid resolution used in our analysis.

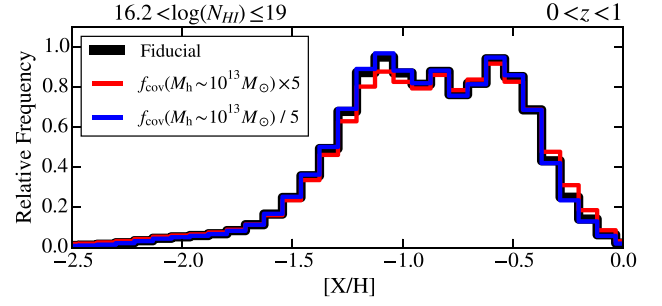




**Figure A2.** Same as Fig. A1 but for the LLS metallicity distribution in the m12i simulation.

simulation are somewhat more time variable than in Fig. 4 in the main text because here we average over only a single quadrant of the simulation, whereas Fig. 4 averages over the entire projected virial area.

The most massive simulated haloes in our sample, m13 and MFz0\_A2, were run at one resolution level lower (gas particle masses  $m_b \approx (3 - 4) \times 10^5 M_\odot$ ) than the rest of our simulation sample. To test whether our results are sensitive to possible resolution effects in these simulations, we artificially varied their LLS covering fractions by a factor of 5 up and down and re-computed the cosmological metallicity distribution. Fig. A3 shows the results.



**Figure A3.** Overall cosmological LLS metallicity distribution recomputed by artificially increasing (red) and decreasing (blue) by a factor of 5 the weights of the most massive haloes (m13 and MFz0\_A2), which are simulated with coarser mass resolution. The good agreement with the fiducial distribution (black) indicates that these contribute relatively little to the cosmological LLS metallicity distribution. The lower resolution of these simulations is thus unlikely to significantly affect our results.

Even if these massive haloes were subject to significant resolution limitations, our main results would be affected only very slightly. This is because haloes in this mass range are rare and contribute negligibly to global LLS statistics.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.